The Appeal of Political Liberalism

Samuel Scheffler


Stable URL:
http://links.jstor.org/sici?sici=0014-1704%28199410%29105%3A1%3C4:ATAOPL%3E2.0.CO%3B2-Q

*Ethics* is currently published by The University of Chicago Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ucpress.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
The Appeal of Political Liberalism*

Samuel Scheffler

The appeal of liberalism derives to a considerable extent from its commitment to tolerating diverse ways of life and schemes of value. Yet this same commitment is also responsible for much of what is puzzling about liberalism. For what is the basis of liberal toleration? One answer rests the case for toleration on a pluralistic understanding of the nature of human value, on a conviction that the realm of value is irreducibly heterogeneous. Diverse ways of life should be tolerated, on this view, because they are routes to the realization of diverse human goods. A very different answer rests the case for toleration on a general skepticism about value, on a conviction that there is no good sense to be made of the idea of objective value or the notion of a good life. On this view, diverse ways of life should be tolerated because there is nothing to the thought that some ways of life are better than others, and so there is no legitimate basis for intolerance.

If the case for liberal toleration rests on some pluralistic thesis about the nature of human values, then both the depth of such toleration and the extent of its appeal seem called into question. For, inevitably, the pluralistic thesis will itself be controversial. Thus, on this interpretation, liberalism's professed toleration of differing conceptions of value turns out to depend on a more fundamental commitment to a particular conception of value, a conception which will be uncongenial or even abhorrent to some of the very evaluative outlooks that liberalism purports to tolerate, and which will not, therefore, serve to recommend liberal institutions to people who share those outlooks. Much the same will be true, it seems, if toleration is seen as the outgrowth of skepticism rather

* This essay is a slightly revised version of a paper presented at the Third Annual Philosophy Conference at the University of California, Riverside, on May 1, 1993. I am indebted to those present on that occasion for helpful discussion, and especially to David Gill, who raised the question that prompted most of the revisions in the present version. I am also indebted to John Rawls for illuminating discussion of his work over a period of many years.

*Ethics* 105 (October 1994): 4–22
© 1994 by The University of Chicago. All rights reserved. 0014-1704/95/0501-0001$01.00
than pluralism about value. For skepticism no less than pluralism represents a controversial understanding of the nature of values.

The most obvious way of defending liberal toleration, apart from the two just mentioned, is as a modus vivendi, a strategic compromise among contending social groups, none of whom is in a position to impose its preferred way of life on the others without intolerable cost, and each of whom therefore accepts a policy of mutual tolerance as the best that it can hope to achieve under the circumstances. That it will sometimes be possible to make a strong case for liberal toleration on pragmatic grounds of this sort seems hard to deny. Historically, the role played by such considerations has often been crucial. Yet a defense of toleration that rests entirely on pragmatic grounds seems unable to account for the moral appeal of the idea of tolerance, and, in any case, it retains its force only so long as the necessary balance of power in society is preserved. If one group gains enough strength that a policy of intolerance comes to seem tempting, the pragmatic argument provides no reason to resist the temptation.

The need to understand the basis of liberal toleration has taken on a renewed urgency at this historical moment. The liberal societies of the west are beset by a host of social problems whose source, according to many critics, lies ultimately in a culture of individualism and a breakdown of communal values for which liberal thought itself is responsible. While the tendency of this criticism is to suggest a need for greater social unity and cohesion, the very idea of a liberal society as a single national community with a common culture is also under attack, as such societies, with their increasingly diversified populations, struggle to come to terms with their own histories of exclusion and to accommodate the claims of "multiculturalism." In this context, it is more important than ever to arrive at a clear understanding of the basis of liberal toleration. More generally, liberalism needs to understand how it is to conceive of its relations to the diverse ways of life and forms of culture that characterize modern societies. The importance of this project of liberal self-understanding, as I shall refer to it, is only intensified by the need for liberal societies to orient themselves both in relation to those other societies, in Eastern Europe and elsewhere, that are struggling to establish liberal institutions for the first time, and in relation to societies like Iran, whose fundamentalist character presents a radical challenge to liberal thought.

John Rawls's book *Political Liberalism* constitutes a major contribution to this project of liberal self-understanding. Rawls offers us a new way of thinking about the basis of liberal toleration. His "political

---

liberalism" addresses itself to the following question: "How is it possible for there to exist over time a just and stable society of free and equal citizens who still remain profoundly divided by reasonable religious, philosophical, and moral doctrines?" (p. 47). The core of political liberalism's answer to this question is that for such a society to be possible, its basic structure must be "effectively regulated by a political conception of justice that is the focus of an overlapping consensus of at least the reasonable comprehensive doctrines affirmed by its citizens" (p. 48). To understand this answer, obviously, we need to understand what is meant by notions like a "political conception of justice" and an "overlapping consensus."

To begin, we know from A Theory of Justice that a conception of justice is a "set of principles for assigning basic rights and duties and for determining... the proper distribution of the benefits and burdens of social cooperation." A political conception of justice, Rawls now tells us, is one that has "three characteristic features" (p. 11). First, it is a moral conception, but "a moral conception worked out for a specific kind of subject" (p. 11), namely, for the "basic structure" of society, by which Rawls means a society's main social, political, and economic institutions. Second, a political conception of justice "is presented as a freestanding view" (p. 12). This means that "it is neither presented as, nor as derived from, [a comprehensive moral] doctrine applied to the basic structure of society, as if this structure were simply another subject to which that doctrine applied" (p. 12). Rather, a political conception "is a module, an essential constituent part, that fits into and can be supported by various reasonable comprehensive doctrines that endure in the society regulated by it" (p. 12). Third, the content of a political conception of justice "is expressed in terms of certain fundamental ideas seen as implicit in the public political culture of a democratic society" (p. 13). The suggestion here is that despite the diversity of people's comprehensive moral doctrines, there may be certain fundamental ideas implicit in a society's culture and institutions that command widespread agreement. A political conception of justice is one that is developed on the basis of this common ground.

If a political conception of justice can indeed be derived solely from ideas that are latent in the public political culture, then it may become the object of an "overlapping consensus." That is, it may be accepted by people who affirm very different comprehensive moral doctrines, because all of these people accept, albeit for different reasons, the fundamental ideas that function, in effect, as premises in the argument for the conception of justice. In an overlapping consensus,

Rawls writes, “the reasonable [comprehensive] doctrines endorse the political conception, each from its own point of view. Social unity is based on a consensus on the political conception; and stability is possible when the doctrines making up the consensus are affirmed by society’s politically active citizens and the requirements of justice are not too much in conflict with citizens’ essential interests as formed and encouraged by their social arrangements” (p. 134).

Not only will people’s reasons for accepting the political conception vary, depending on which comprehensive moral doctrine they affirm, but in addition, different comprehensive moral doctrines may stand in different relations to the political conception. For example, some comprehensive doctrines may provide the political conception with a deductive basis, while others may provide reasons of other kinds for accepting it. Rawls says that “the point to stress here is that . . . citizens individually decide for themselves in what way the public political conception all affirm is related to their own more comprehensive views” (p. 38). What matters, he writes, is that “citizens themselves, within the exercise of their liberty of thought and conscience, and looking to their comprehensive doctrines, view the political conception as derived from, or congruent with, or at least not in conflict with, their other values” (p. 11).

These ideas enable Rawls to offer a model of liberal pluralism and toleration that represents an alternative to the three I canvassed at the outset. Like the model of liberalism as a modus vivendi, Rawls’s model treats with utmost seriousness the deep differences in people’s values and ways of life, and the potential for deadly conflict arising out of those differences. Yet, as he insists, the overlapping consensus that he envisions is not itself a mere modus vivendi. For, first, the object of the consensus is not just a set of institutional arrangements but is, rather, a moral conception, a conception of justice. And, second, all those who participate in the consensus accept that conception for moral reasons of one sort or another. As Rawls says, “All those who affirm the political conception start from within their own comprehensive view and draw on the religious, philosophical, and moral grounds it provides. The fact that people affirm the same political conception on those grounds does not make their affirming it any less religious, philosophical, or moral, as the case may be, since the grounds sincerely held determine the nature of their affirmation” (pp. 147–48). Because of these two differences between an overlapping consensus and a modus vivendi, there will also, Rawls says, be a third difference. An overlapping consensus will enjoy greater stability than a modus vivendi, because the commitment to the political conception of those who participate in the consensus is based on moral considerations rather than calculations of self or group interest and hence is not liable to be undermined by changes in the balance of power within the
society. One’s moral reasons for accepting the political conception are not weakened by an increase in one’s power that would make it easier to press for a more advantageous set of institutional arrangements.

In a sense, then, Rawls's explanation of the basis of liberal toleration, if successful, may manage to combine the advantages of an appeal to a modus vivendi with those of an appeal to a pluralistic conception of value, while avoiding the pitfalls of either. Like the explanation in terms of a modus vivendi and unlike the one in terms of a pluralistic conception of value, Rawls combines respect for the facts of disagreement and diversity with a reluctance to rely on any controversial moral premises. However, like the explanation in terms of a pluralistic conception of value and unlike the one in terms of a modus vivendi, Rawls represents citizens as having moral reasons for their allegiance to the structures of toleration and the institutions of the liberal society. His aim, in other words, is to provide liberal institutions with a basis in moral reasons, without himself presupposing any controversial and contentious moral outlook. This aspiration is descended from Rawls's earlier aspiration, in A Theory of Justice, to derive a substantive conception of justice from the set of putatively "weak and widely shared" conditions and constraints that make up the original position. If Rawls can succeed in achieving his newer aim, then the earlier one is also vindicated, at least in part. For the notion of an overlapping consensus on fundamental ideas implicit in the public political culture supplies new content for the claim that the original position—construed now as a device for representing certain of those fundamental ideas—is made up of weak and widely shared assumptions.

Thus, from the standpoint of the project of liberal self-understanding, the potential advantages of "political liberalism" seem clear. Ultimately, however, the appeal of the view must depend, at least in part, on whether an overlapping consensus on something like Rawls's political conception of justice is a realistic possibility. In chapter 4 of Political Liberalism, Rawls offers what he calls "a model case" of such a consensus:

It contains three views: one affirms the political conception because its religious doctrine and account of free faith lead to a principle of toleration and underwrite the fundamental liberties of a constitutional regime; while the second view affirms the political conception on the basis of a comprehensive liberal moral doctrine such as those of Kant or Mill. The third . . . is not systematically unified: besides the political values formulated by a freestanding political conception of justice, it includes a large family of nonpolitical values. It is a pluralist view, let us say, since each subpart of this family has its own account based on ideas drawn from within it, leaving all values to be balanced against one another, either in groups or singly, in particular kinds of cases. [P. 145]
By itself, this model case does little to encourage the thought that an overlapping consensus on Rawlsian principles could actually be achieved in a society like ours. For the three views it contains are not fully representative of the diverse schemes of value one finds in modern liberal societies; instead, all three appear to be drawn from the same relatively narrow portion of the broad spectrum of evaluative conviction. Although it seems plausible enough that these three views might converge on a liberal conception of justice, that tells us little about the prospects for a more inclusive consensus.

Later in chapter 4, Rawls presents a modified version of this model case which also includes classical utilitarianism. "This utilitarianism," he writes, "supports the political conception for such reasons as our limited knowledge of social institutions generally and on our knowledge about ongoing circumstances. It stresses further the bounds on complexity of legal and institutional rules as well as the simplicity necessary in guidelines for public reason. These and other reasons may lead the utilitarian to think a political conception of justice liberal in content a satisfactory, perhaps even the best, workable approximation to what the principle of utility, all things tallied up, would require" (p. 170).

Although the inclusion of classical utilitarianism serves to broaden the range of outlooks contained in the model consensus, and although its explicit purpose is just to illustrate the point that the different comprehensive doctrines represented in an overlapping consensus may stand in different relations to the political conception on which they converge, it is in some ways a puzzling addition. Rawls emphasizes that an overlapping consensus is a consensus not just on principles of justice but also on the fundamental ideas implicit in the public political culture from which those principles are derivable (p. 149). Indeed, what makes an overlapping consensus on a political conception of justice possible is precisely the fact that the political conception is developed from shared ideas. Accordingly, the original position is now to be construed as modeling certain of those shared ideas, and Rawls's arguments to the effect that his principles would be chosen in the original position are to be interpreted as beginning from those ideas. Yet many of these arguments are explicitly directed against utilitarianism. If utilitarianism is said to be included in the overlapping consensus on Rawls's two principles, then are we to imagine that utilitarians endorse Rawls's arguments for the rejection of utilitarianism even as they continue to affirm that view? This seems incoherent.

Moreover, even the fundamental ideas from which the arguments for the two principles proceed, and which the original position helps to model, are ideas which, according to Rawls himself, utilitarianism does not accept. For example, the "fundamental organizing idea" to which Rawls appeals is "that of society as a fair system of social coopera-
tion between free and equal persons viewed as fully cooperating members of society over a complete life” (p. 9). Crucial to this idea of fair cooperation is said to be a notion of reciprocity, according to which “all who are engaged in cooperation and who do their part as the rules and procedures require, are to benefit in an appropriate way as assessed by a suitable benchmark of comparison” (p. 16). Yet in A Theory of Justice Rawls tells us that “the principle of utility is incompatible with the conception of social cooperation among equals for mutual advantage. It appears to be inconsistent with the idea of reciprocity implicit in the notion of a well-ordered society.”

Elaborating on this point, he writes: “Implicit in the contrasts between classical utilitarianism and justice as fairness is a difference in the underlying conceptions of society. In the one we think of a well-ordered society as a scheme of cooperation for reciprocal advantage regulated by principles which persons would choose in an initial situation that is fair, in the other as the efficient administration of social resources to maximize the satisfaction of the system of desire constructed by the impartial spectator from the many individual systems of desires accepted as given.”

If this is right, then it really is quite unclear how utilitarianism can be included in an overlapping consensus on Rawls’s principles of justice, since it rejects the fundamental ideas that serve as premises in the arguments for those principles, as well as the arguments themselves.

Of course, as Rawls says, a utilitarian might conclude that, “given normal social conditions” (p. 171), Rawls’s two principles represent “a satisfactory, perhaps even the best, workable approximation to what the principle of utility, all things tallied up, would require” (p. 170). The question is whether such a conclusion would be sufficient for utilitarianism to be included in an overlapping consensus, if it is also committed to rejecting the fundamental ideas from which Rawls’s principles are derived as well as the arguments used to derive them. Moreover, the precise content and spirit of the imagined utilitarian conclusion are less clear than they may initially appear. We may recall that in A Theory of Justice Rawls considers two possible attitudes a utilitarian might take toward apparently nonutilitarian principles of justice. On the one hand, some utilitarians may say that “common sense precepts of justice and notions of natural right have but a subordinate validity as secondary rules; they arise from the fact that under the conditions of civilized society there is great social utility in following them for the most part and in permitting violations of them only under exceptional circumstances.”

Rawls’s project in A Theory of Justice is to a large extent

4. Ibid., p. 33.
5. Ibid., p. 28.
motivated by the conviction that this attitude does not accord sufficient primacy to the principles of justice. On the other hand, a utilitarian might, Rawls suggests, argue that utility will actually be maximized if Rawls's two principles are "publicly affirmed and realized as the basis of the social structure." However, Rawls maintains, such an argument would be tantamount to conceding the inadequacy of utilitarianism. For the publicity condition means that utilitarianism must be "defined" as "the view that the principle of utility is the correct principle for society's public conception of justice." In other words, Rawls says, "what we want to know is which conception of justice characterizes our considered judgments in reflective equilibrium and best serves as the public moral basis of society. Unless one maintains that this conception is given by the principle of utility, one is not a utilitarian." The upshot is that of the two possible attitudes that a utilitarian might take toward nonutilitarian principles of justice, the first represents too weak a form of endorsement, while the second is tantamount to the abandonment of utilitarianism. Thus neither attitude provides a model of how a committed utilitarian might genuinely affirm Rawls's principles as part of an overlapping consensus, and this only reinforces the doubts already mentioned about whether utilitarianism could indeed be included in such a consensus.

Even if it could, the question of how widespread a consensus on Rawls's principles of justice might realistically be achieved would remain. For a stable overlapping consensus in a society like ours would require the inclusion not only of fully articulated philosophical theories like utilitarianism but also of the many moral, religious, and broadly philosophical outlooks actually endorsed by citizens in our society. A stable overlapping consensus, one might say, must be a consensus of citizens, not of theories. Rawls recognizes this, and in chapter 4 of Political Liberalism, he tries to turn the point to his advantage. He argues, in effect, that in judging the prospects of an overlapping consensus on his conception of justice, it is a mistake to focus exclusively on committed partisans of fully articulated philosophical theories with clearly recognized implications for the justice of the basic structure of society. For most citizens do not hold any fully articulated and fully comprehensive moral doctrine. What they may well have instead is a loose network of values and convictions of varying degrees of generality, whose implications concerning Rawls's principles may or may not be clear to them. However, many of these people, having been raised in a liberal society, will have internalized the fundamental

6. Ibid., p. 181.
7. Ibid., p. 182.
8. Ibid.
idea of society as a fair system of cooperation on which Rawls's political conception rests and will have developed a strong allegiance to broadly liberal values and institutions. Indeed, rather than accepting liberal ideas as the consequence of their comprehensive moral doctrine, they may instead allow their commitment to a liberal order to shape the rest of their values. For example, they may find themselves shaping their understanding of their faith to render it consistent with such an order—perhaps emphasizing those strands in their tradition that lend support to tolerance and downplaying those that conflict with it. If this is correct, and if Rawls can in fact elaborate his conception of justice solely on the basis of ideas implicit in the public political culture of a liberal society, then such people may have strong reasons to affirm that conception. As Rawls puts it, "the comprehensive doctrines of most people are not fully comprehensive, and this allows scope for the development of an independent allegiance to the political conception that helps to bring about a consensus" (p. 168).

Although this argument has some force, there is an important ambiguity in Rawls's account of the sort of overlapping consensus he envisions, and this ambiguity affects the plausibility of his position. Recall that in an overlapping consensus, what people are said to affirm is a political conception of justice. And "a distinguishing feature of a political conception is that it is presented as freestanding" (p. 12), that is, as independent of any comprehensive moral doctrine. Now this suggests two puzzles. First, by whom must a conception be presented as freestanding in order to count as political? One possibility is that Rawls's presentation of his own conception is authoritative, so that that conception is political because he presents it as freestanding. But then, in order to know whether a given conception of justice is political we must be able to identify its authoritative presentation. What, for example, is the authoritative presentation of utilitarianism? Another possibility is that conceptions of justice are to be individuated partly by features of their presentation. Thus if you argue for Rawls's conception of justice by appealing to your comprehensive moral doctrine and I argue for it as a freestanding view, we are actually arguing for two different conceptions of justice, one political and the other not. But this seems needlessly confusing and potentially misleading, and in any case, it is a peculiar position to take if one is attempting to emphasize the possibility of consensus—to explain how people with divergent moral outlooks may nevertheless converge on a particular conception of justice. Still another interpretation is suggested when Rawls remarks at one point that what is characteristic of a political conception is only "that it can be presented without saying, or knowing, or hazarding a conjecture about, what . . . [comprehensive] doctrines it may belong to, or be supported by" (pp. 12–13; emphasis added). This leaves open the possibility that all conceptions of justice may be political, if
all of them can be “presented as” freestanding. (How plausible would the presentations have to be?)

None of these three options seems to me especially attractive. It might be less confusing and more illuminating to use the adjective ‘political’ to describe arguments for conceptions of justice rather than the conceptions themselves. A political argument for a conception of justice would be one that appealed to ideas implicit in the public political culture, whereas a nonpolitical argument, say, would be one that appealed to a comprehensive moral doctrine. Thus one and the same conception of justice might in principle be supported by arguments of either type. Rawls might then be interpreted as asserting not that his conception of justice is a political conception but, rather, that his arguments for that conception are political arguments. And political liberalism might be construed as comprising two theses: first, that given the pluralistic character of modern societies, a just and stable order is possible only if the basic structure of society is effectively regulated by a conception of justice that is the focus of an overlapping consensus and, second, that the possibility of an overlapping consensus on a particular conception of justice is signaled by the availability of persuasive political arguments for that conception.

There is, however, a second and deeper puzzle underlying the one just mentioned, and it concerns the attitude that participants in an overlapping consensus are expected to have toward the “political conception of justice” on which they converge. In particular, the question is whether participation in such a consensus requires that one regard the conception of justice as a “political conception”: that is, as a “freestanding” conception whose content “is expressed in terms of certain fundamental ideas seen as implicit in the public political culture of a democratic society.” We know, in other words, that the participants in an overlapping consensus accept certain principles of justice as well as certain fundamental ideas implicit in the public political culture from which those principles are derivable. The question is whether they themselves also think of the principles as being expressed in terms of what they take to be ideas implicit in the public political culture, and as capable of being derived independently of any particular comprehensive moral doctrine. On the face of it, it would appear unwise to make this a requirement for participation in an overlapping consensus. For the more things that people must believe in order to be included in such a consensus, the more difficult it will be for a consensus actually to be achieved. In other words, if participation in the consensus requires affirmation not only of a particular set of principles of justice but also of certain metatheses about the status of those principles, then, other things equal, one would expect the consensus to include fewer people. Furthermore, Rawls's whole reason for drawing our attention to the possibility of an overlapping consen-
sus is to suggest that people may affirm the same principles of justice even though they view these principles in very different ways, depending on which comprehensive moral doctrine they accept. Thus any requirement that the participants in an overlapping consensus must view the conception of justice as political would appear to be incongruous with the motivation for introducing the idea of such a consensus in the first place.

Suppose, for example, that someone—call her Jane—accepts Rawls's two principles of justice because she believes them to be implied by her comprehensive moral doctrine, which we may suppose to be some version of what Rawls calls "comprehensive liberalism." Jane does not, let us assume, think of the two principles as constituting a freestanding conception of justice, nor does she characterize them as such when she is explaining or defending them to others. Rather, she presents them to others as she herself conceives of them, namely, as derived from a certain comprehensive moral outlook. And although she recognizes that the two principles may be seen as giving expression to the idea of society as a fair system of cooperation, it does not occur to her to think of that idea as implicit in the public political culture; still less does she see the fact that it is implicit in the culture as crucial to its justificatory role. The importance to her of the idea of society as a fair system of cooperation derives instead from its relation to the ideal of "ethical autonomy" (pp. 77–78) that is at the core of her comprehensive doctrine. Yet Jane is, we may suppose, a sincere and indeed committed advocate of the two principles of justice and, as such, she is of course a staunch defender of the right of other people to affirm and defend comprehensive moral doctrines other than her own. If, by virtue of her failure to conceive of the two principles as amounting to a political conception of justice, a person like Jane does not qualify for inclusion in an overlapping consensus, then surely the prospects of achieving a widespread consensus are open to serious doubt. Moreover, to exclude Jane would seem inconsistent with the motivation for introducing the idea of an overlapping consensus to begin with.

Nevertheless, Rawls does at times appear to require that the participants in an overlapping consensus regard the conception of justice as political. He writes, for example, that "citizens are to conduct their fundamental discussions within the framework of what each regards as a political conception of justice based on values that the others can reasonably be expected to endorse and each is, in good faith, prepared to defend that conception so understood" (p. 226). Rawls makes this claim in the context of his discussion of "the idea of public reason." This idea arises because, Rawls believes, "it is essential that a liberal political conception include, besides its principles of justice, guidelines of inquiry that specify ways of reasoning and criteria for the kinds of
information relevant for political questions” (p. 223). He therefore maintains that “the parties in the original position, in adopting principles of justice for the basic structure, must also adopt guidelines and criteria of public reason for applying those norms” (p. 225). These guidelines and criteria are to specify the modes of reasoning that may be used and the types of considerations that may be appealed to in discussing and resolving political questions in a society regulated by the principles of justice. They impose constraints on acceptable forms of political argument, constraints that Rawls refers to as “the limits of public reason.” These limits hold not only for public officials but also for “citizens when they engage in political advocacy in the public forum, and thus for members of political parties and for candidates in their campaigns and for other groups who support them” (p. 215). Indeed, the limits govern “how citizens are to vote” (p. 215), at least when “the most fundamental political questions” are at stake (p. 216).

Among the most important of the constraints imposed by the limits of public reason is that “in discussing constitutional essentials and matters of basic justice we are not to appeal to comprehensive religious and philosophical doctrines—to what we as individuals or members of associations see as the whole truth” (pp. 224–25). In other words, the limits of public reason tell us that we are “to conduct our fundamental discussions in terms of what we regard as a political conception. We should sincerely think that our view of the matter is based on political values everyone can reasonably be expected to endorse” (p. 241).

Rawls’s discussion of public reason raises a variety of questions. For one thing, it seems puzzling to suggest that the parties in the original position adopt guidelines that specify, among other things, that we are “to conduct our fundamental discussions in terms of what we regard as a political conception” of justice. For the idea of a political conception of justice is not one that is obviously available to the parties in the original position. What makes Rawls’s conception political is the fact that it is elaborated on the basis of fundamental ideas that are implicit in the public political culture. The function of the parties in the original position is to help model certain of those fundamental ideas, not to endorse them or to argue from them to something the parties regard as a political conception. In other words, the parties’ sole concern is to choose a conception of justice that will maximize their share of primary social goods. They do not themselves decide that they want a political conception of justice based on ideas implicit in the public political culture; rather, the fact that they model such ideas is that we have reason to regard the principles they select as having the status of a political conception.

More generally, the idea that the parties in the original position adopt “guidelines of inquiry” which include “principles of reasoning and rules of evidence” (p. 224) requires further explanation. Rawls
says that the "argument for those guidelines . . . is much the same as, and as strong as, the argument for the principles of justice themselves" (p. 225). Perhaps so; but in the case of the guidelines, the array of options from which the parties choose and the basis of their choice require further elaboration. Moreover, there is the following difference between the choice of such guidelines and the choice of principles of justice. In order to make their choice among rival conceptions of justice, the parties do not themselves need to employ any particular conception of justice. If they did, the justificatory force of their choice would be open to serious doubt. In choosing among candidate principles of reasoning and methods of inquiry, however, the parties must already be employing certain methods of reasoning and modes of inquiry, which are given by stipulation as part of the original position construction. Thus, in the absence of some further explanation of the relation between the methods of inquiry the parties employ and the methods of inquiry they adopt, there is a danger that any choice they make will appear question begging.

For present purposes, however, what is most striking about Rawls's discussion of public reason is how emphatically it serves to reinforce the concern that by requiring the participants in an overlapping consensus to conceive of the conception of justice they endorse as political, Rawls risks undermining the plausibility of the idea that an overlapping consensus might actually be achieved. The principle that ordinary citizens, when engaged in political advocacy and even when voting, must appeal only to what they regard as a political conception of justice and never to their own comprehensive moral doctrines is an extraordinarily strong one. Accordingly, it seems much harder to envision a wide variety of comprehensive doctrines converging on this principle of public reason in addition to the two principles of justice than it does to envision them converging on the principles of justice alone. Indeed, the plausibility of this putative principle of public reason is open to serious challenge. For it does not seem difficult to think of instances in our own society in which people have appealed to comprehensive moral doctrines in ways that many would regard as appropriate, or at least not inappropriate. Toward the end of his discussion of public reason, Rawls himself considers the examples of the abolitionists and of Martin Luther King, Jr., and in the face of these examples he modifies his principle to allow that citizens may in certain circumstances appeal to their comprehensive doctrines if this is necessary to strengthen the ideal of public reason itself. But there are many other examples that could be cited in addition to the two that Rawls mentions: such as the religiously inspired opposition of Quakers and others to the war in Vietnam; the religiously motivated opposition of many people to capital punishment; the central role traditionally played by black churches in the political life of the African-American
community; the opposition by certain religious denominations to United States policies in Central America during the 1980s and the associated movement to provide sanctuary in churches for Central American refugees; and the religiously based advocacy of policies to eliminate homelessness and poverty. Note, moreover, that all of the examples I have cited are instances in which comprehensive religious doctrines have been drawn on in the context of public political advocacy, in ways generally congenial to the political left. If we broaden our horizons to include nonreligious appeals to moral principles, to include citizens’ reasons for voting as well as the content of their public advocacy, and to include positions taken across a wider range of the political spectrum, then the number of examples will only multiply. The idea that in all such examples the reliance on a comprehensive moral or religious doctrine must either be necessary to strengthen the ideal of public reason or else unjustified seems highly questionable. And in any event, the availability of so many examples does nothing to enhance the plausibility of the idea that an overlapping consensus might actually converge not only on Rawls’s two principles of justice but also on the guidelines of public reason as he describes them.

Given the contentiousness of some of Rawls’s claims about the limits of public reason, and given that their effect is also to make the achievement of an overlapping consensus seem less likely, it is natural to wonder how important it is for Rawls’s purposes to insist on those claims. In an attempt to ascertain this, let us imagine a society in which an overlapping consensus on Rawls’s two principles of justice has actually been achieved and in which a constitution consistent with those two principles has been adopted. Let us further suppose, however, that citizens in this society view the status of the two principles in different ways, depending on their comprehensive moral doctrines. Thus there is no consensus on the proposition that the two principles of justice constitute a political conception, still less on the proposition that citizens are to conduct their discussion of constitutional essentials and basic justice in terms of what they regard as a political conception. Despite the absence of consensus on these two points, the constitution of the society guarantees the basic rights and liberties of all citizens, in accordance with the principles of justice, and this excludes any attempt to use the coercive power of the state to enforce a particular comprehensive doctrine. Similarly, any attempt to deny basic justice to some people in the name of a particular comprehensive doctrine is ruled out. Moreover, since all citizens, by hypothesis, affirm that the two principles are to regulate the basic structure of their society, they all regard those principles as the “final court of appeal for ordering the conflicting claims of moral persons.” 9 Thus it is to the two

9. Ibid., p. 135.
principles, rather than their own comprehensive moral doctrines, that citizens naturally appeal when discussing constitutional essentials and matters of basic justice. In addition, citizens are aware of the existence both of widespread agreement on the principles of justice and of widespread disagreement in people's comprehensive moral doctrines. This awareness, when coupled with the commitment to mutual respect that is implied by citizens' common affirmation of the two principles, gives rise to an ethos of restraint, a reluctance on the part of many citizens to appeal in the public arena to their own comprehensive moral doctrines. After all, the fact that they are in agreement about the principles to regulate the basic structure of their society is what matters for fundamental political purposes, and, we may suppose, they have no need and little desire to alienate those they respect by insisting on divisive moral or religious claims.

The upshot of this discussion is that some limits of public reason may be a consequence of, and still others may be encouraged by, the existence of an overlapping consensus on the two principles alone, without any need for citizens to conceive of those principles as a political conception of justice or to agree that they must conduct their discussion of fundamental political questions in terms of what they regard as a political conception. If this is correct, and if the requirement that citizens view the conception of justice as political has the disadvantages I have said it does, then Rawls may have little reason to insist on that requirement.

The point I have been trying to make may be put another way. One of Rawls's aims in *Political Liberalism*, if I understand him correctly, is to respond to those critics of his earlier work who charged that while purporting to offer a neutral and universal justification for his principles of justice, it rested tacitly but unmistakably on a liberal conception of value. In response, Rawls denies that his theory aspires to universal validity but also that it rests on any comprehensive conception of value. Instead, he argues, his theory is addressed to societies of a certain type at a particular historical moment. These societies have a tradition of democratic thought and constitutional interpretation but there exist within them deep disagreements about fundamental political questions and also a wide diversity of comprehensive moral and religious doctrines. Given these historical facts, there is a real question about how a shared and workable conception of justice for these societies can be arrived at. It is to this question that Rawls's theory is addressed, and his answer involves not the endorsement of a particular comprehensive doctrine but, rather, an attempt to bypass the disagreements among such doctrines. What he tries to do, as we know, is to identify certain bases of agreement that are implicit in the public political culture and which therefore represent common ground among the citizens of democratic societies. He then attempts to use
these “fixed points” as premises in an argument for a conception of justice to which all or nearly all may be able to agree. This is what Rawls is trying to do: to argue from views that are widely shared in our culture to a definite conclusion about justice, thereby offering a conception of justice that may command widespread agreement despite the pluralism and disagreement that characterize our society. If his argument is successful, then the conception he identifies may serve as the object of an overlapping consensus among people who have different comprehensive moral doctrines and whose affirmation of the theory is therefore based ultimately on different reasons. In order to be successful, it is important that Rawls’s argument be a “political” one: that is, that it rely as much as possible on shared ideas and avoid reliance on any comprehensive moral doctrine. However, if what I have been urging is correct, it is not important that the citizens who are included in the overlapping consensus should themselves think of Rawls’s conception as political. Any requirement that they do so would make an overlapping consensus more difficult to achieve, would add little that is plausible to an adequate account of public reason, and would mandate a degree of metaethical uniformity that is incongruous with the motivation for introducing the idea of an overlapping consensus in the first place. For all of these reasons, such a requirement would weaken rather than strengthen Rawls’s reply to his critics. Or so it seems to me.

This conclusion is not altered, incidentally, if we take into account the added complication that Rawls characterizes his conception of justice not only as political but also as constructivist. It is constructivist because its principles “may be represented as the outcome of a certain procedure of construction” (pp. 89–90). That is, the principles are constructed from ideas of society and the person via the device of the original position. A constructivist conception, Rawls says, neither asserts nor denies that its principles are true; it “does without the concept of truth” altogether (p. 94). Political constructivism, which is “part of” political liberalism, asserts instead that the constructivist conception is “reasonable for a constitutional regime” (p. 126). Although many citizens “may want to give the political conception a metaphysical foundation as part of their own comprehensive doctrine” (p. 126), political constructivism refrains from doing this and restricts itself to the claim that the conception is reasonable. “The advantage of staying within the reasonable,” Rawls writes, “is that there can be but one true comprehensive doctrine, though . . . many reasonable ones. Once we accept the fact that reasonable pluralism is a permanent condition of public culture under free institutions, the idea of the reasonable is more suitable as part of the basis of public justification for a constitutional regime than the idea of moral truth. Holding a political conception as true, and for that reason alone the one suitable
basis of public reason, is exclusive, even sectarian, and so likely to foster political division” (p. 129). The question, however, is who exactly is supposed to believe that “the idea of the reasonable is more suitable as part of the basis of public justification for a constitutional regime than the idea of moral truth.” It is easy to see that there may be an advantage to Rawls in offering a purely constructivist argument for his principles—in asserting, that is, not that the principles are true but merely that their source in ideas implicit in the public political culture enables them to provide a reasonable basis of public justification. For if he can give people a reason to affirm his principles whether or not they accept the idea of moral truth, he maximizes the appeal of those principles and enhances the prospects of an overlapping consensus. He makes it possible, in other words, for people to agree on the principles even if they disagree about the metaphysical status of those principles. But if citizens are also required to agree that “the idea of the reasonable is more suitable as part of the basis of public justification . . . than the idea of moral truth,” then Rawls risks squandering these very gains, for he is then insisting on a partial but nevertheless significant degree of uniformity in the way people regard his principles.

Up to this point, I have been exploring the appeal of political liberalism by examining the prospects for an overlapping consensus on a political conception of justice in a modern democratic society like our own. Before concluding, I want to take up one other issue that bears on the appeal of political liberalism, and this concerns the relevance of this form of liberalism for those societies whose traditions are not liberal or democratic. Thus I want to turn from political liberalism’s domestic policy, as it were, to its foreign policy. As we have seen, Rawls’s work as he now presents it is addressed to modern democratic societies at a certain historical moment. His political liberalism seeks to establish a liberal conception of justice on the basis of ideas that are implicit in the public political culture of such societies. If, however, political liberalism appeals in this way to the public political culture of democratic societies, then the justification of liberal principles and institutions that it offers appears to presuppose a society in which liberal values are already well entrenched. It is not clear that political liberalism provides any reason for establishing liberal institutions in societies that do not already have liberal traditions. Thus, to put it another way, it is not clear that political liberalism could ever provide the original justification for a society’s liberal institutions. Liberal traditions and institutions must, it seems, precede political liberalism: that is, they must already exist in order to create the conditions that make possible the sort of justification offered by political liberalism. This is not merely a historical problem; it raises doubts about whether political liberalism has anything to offer to those aspiring democracies in East-
ern Europe and elsewhere that have no liberal traditions and whose public political cultures therefore lack the implicit ideas to which Rawls appeals. If in fact political liberalism has nothing to offer such societies, then its defense of liberal principles and institutions may seem intolerably weak. In renouncing any universalistic ambitions, Rawls may now seem to have gone too far in the other direction and to have produced a version of liberalism that is so historically specific and so dependent on a prior history of liberal institutions as to be of little relevance in those situations where the justification of liberalism matters most: that is, where liberalism is confronted by, and must engage with, societies whose traditions and practices are not liberal.

In a recent essay titled “The Law of Peoples,” Rawls makes clear that his defense of liberalism does not in fact aspire to the kind of universality that some liberals would prefer.\(^{10}\) In that lecture, he attempts to develop what he calls a “liberal law of peoples,” which applies even to societies that are not themselves liberal. Yet he also makes it clear that this law does not require nonliberal societies to become liberal. On the contrary, Rawls argues that “not all regimes can be reasonably required to be liberal,” and that a society that is not liberal may nevertheless be “well ordered and just.”\(^{11}\) Rawls’s development of the law of peoples raises many questions that I cannot pursue here. For present purposes, however, what is noteworthy is that it appears to confirm the suspicion that Rawls’s justification of liberal institutions is limited in certain striking ways. Political liberalism makes no general claim about the superiority of liberal over nonliberal societies, nor does it provide arguments as to why heretofore nonliberal societies should become liberal. Rawls seems to regard this reticence as a virtue, in part because he is sensitive to the charge that more ambitious claims on behalf of liberalism represent a form of western ethnocentrism, and because he believes that it is not necessary to make such claims in order to develop a law of peoples that applies to all nations and which specifies certain minimum standards of acceptable conduct.

Yet, it may be objected, political liberalism does not merely refrain from asserting that all societies should become liberal; what is more disturbing is that it cannot give any nonliberal society a reason why it should become liberal. It simply has nothing to offer to those societies that may be attempting, in the face of considerable opposition and without the benefit of any significant democratic tradition, to develop liberal institutions for the first time and which look naturally to the liberal philosophers of the democratic west for what is literally “moral support.”

11. Ibid., pp. 37, 44.
This seems to me a legitimate concern, although Rawls might respond that there is in fact a difference between the reasons why liberal institutions take root in a society for the first time and the justifications for such institutions that become available at later stages. Thus it might be said that liberal institutions take hold, when they do, for a variety of reasons—which may include the need for a modus vivendi, the collapse of alternative institutional schemes, or the desire to emulate the perceived economic success of existing liberal democracies—but which almost certainly do not include a society’s happening to converge from the outset on a particular liberal comprehensive moral doctrine. Moreover, and this may be one of the sobering messages of *Political Liberalism*, liberal institutions founded on the sorts of reasons just mentioned are bound to be precarious until a society has lived under them long enough to develop a tradition of tolerance, a commitment to the virtues of liberal citizenship, and an ethos of reasonableness and fair reciprocity. These things are resources that can help to stabilize liberal institutions by making an overlapping consensus on liberal values and principles possible. But their development takes time and reasonably favorable conditions. It cannot happen overnight at the drop of an argument, even a good one. Once it does happen, however, and this may be the more encouraging message of *Political Liberalism*, then a liberal society may be able to flourish despite the deep disagreements that are bound to persist.

If this response seems insufficiently robust, there is one further thing that might be said consistently with the spirit of Rawls’s book, namely, that if an overlapping consensus on liberal principles can indeed be achieved in modern democracies, then accepting any one of the doctrines included in such a consensus may give one reason to support a liberal scheme. In other words, the distinctive contribution of political liberalism may be to suggest that there are many ways to arrive at liberal principles and that this fact is a source of liberalism’s strength. For if the reasonable comprehensive doctrines that thrive under conditions of freedom all converge on liberal principles, then what may be said to those in societies seeking to establish liberal institutions is that such institutions represent common ground among the various outlooks likely to endure under conditions that encourage the free exercise of human reason. Of course, however, this can be said only if an overlapping consensus is a realistic possibility in modern, pluralistic democracies. This brings us back to the questions raised earlier. Until those sorts of questions are convincingly answered, the suspicion is bound to persist, especially among those who would defend liberalism on the basis of a comprehensive moral doctrine, that Rawls’s position depletes the moral resources of liberalism without managing in exchange to broaden its justificatory appeal.